

Physical and mathematical modeling in experimental papers: achieving robustness of mathematical modeling studies

Vitaly V. Ganusov

Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA

Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA

National Institute for Mathematical and Biological synthesis, Knoxville, TN 37996, USA

February 1, 2016

Abstract

Development of several alternative mathematical models for the biological system in question and discrimination between such models using experimental data is the best way to robust conclusions. Models which challenge existing theories are more valuable than models which support such theories.

Mathematical modeling has been especially strong in physical sciences where theories, based on basic principles, were able to predict energies associated with mass (famous equation $E = mc^2$) and existence of fundamental particles (e.g., positron). There have been many applications of mathematical modeling in biology but because it has been harder to come up with basic principles, predictive successes of mathematical theory-based predictions in biology are more modest than that in physics. With the development of novel experimental methods to collect unique data spanning different levels of organization, from gene sequences and intracellular processes in individual cells to populations of organisms, there has been an impressive rise in the number of mathematical modeling studies in biology. Bioinformatics, commonly defined as a science dealing with analysis of sequencing data (but see (1)), and systems biology, commonly defined as a science dealing with intracellular gene networks, are two most well recognized advances driven by the availability of novel datasets. Several recent reviews focus on the usefulness of mathematical modeling for understanding complex biological systems (2–5). Specifically, Möbius & Laan (3) illustrated how mathematical modeling helps to interpret experimental data using example of molecular motors. While Möbius & Laan (3) do describe some pros and cons of mathematical modeling, there was limited discussion on what makes a good mathematical modeling paper and when mathematical modeling can lead to robust conclusions.

The authors did not cite and did not discuss an important philosophical paper on the use of mathematical modeling in natural sciences, a paper that every mathematical modeler should read, think about, and discuss with his/her peers (6). In contrast with prevailing

view in the mathematical modeling community (e.g., (3)), Oreskes *et al.* (6) argued that for open natural systems (and most if not all systems in biology are open) mathematical models cannot make robust predictions. Moreover, based on philosophical definitions the authors argued that verification of mathematical models is impossible and validation of models, as commonly used in the literature, is nearly impossible (6). They further argued that the use of mathematical models should be rather heuristic, to understand limits of specific mechanisms when models are constrained by experimental data. If most model predictions are not robust (7), how should one proceed with best use of mathematical modeling?

One approach, which was not strongly emphasized by Möbius & Laan (3) is the “method of multiple working hypothesis” by Chamberlin (8) augmented by the “strong inference” of Platt (9). In essence, instead of developing a single mathematical model for the phenomenon in question as it appears to be the norm in the field, one must consider several alternative models and use experimental data as a judge in discriminating between alternative hypotheses (9, 10). In cases when data are not available for model discrimination, analysis of alternative models can inform which experiments may need to be performed to discriminate between alternative models. Furthermore, analysis of multiple alternative models could also highlight limitations on predictions of any individual model, thus, testing robustness of the model predictions which for any given model a priori are expected to not be robust (6, 11). Discrimination between alternative models using experimental data allows to reject some of these models and rejection, rather than confirmation of models is the strongest application of the scientific method (9, 12). Indeed, as Oreskes *et al.* (6) highlighted “models which challenge existing theories are more valuable than models which confirm them.” Confirmation of simplest possible models is reasonable but showing that simple models do not work is clearly more illuminating highlighting our gaps in understanding the system in question (13).

As mathematical modeling is taking an integral part of biology it is time that education and training of mathematical modelers is moving beyond development and analysis of single models. Rather, development of multiple mathematical models which are tested against experimental data and finding which mathematical models are not able explain experimental data needs to become a norm for the mathematical modeling publications in the 21st century.

References

1. Hogeweg, P. 2011 The roots of bioinformatics in theoretical biology. *PLoS Comput Biol*, **7**(3), e1002021. doi:10.1371/journal.pcbi.1002021.
2. Servedio, M. R., Brandvain, Y., Dhole, S., Fitzpatrick, C. L., Goldberg, E. E., Stern, C. A., Cleve, J. V. & Yeh, D. J. 2014 Not just a theory—the utility of mathematical models in evolutionary biology. *PLoS Biol*, **12**(12), e1002017. doi:10.1371/journal.pbio.1002017.
3. Möbius, W. & Laan, L. 2015 Physical and Mathematical Modeling in Experimental Papers. *Cell*, **163**(7), 1577–1583. doi:10.1016/j.cell.2015.12.006.
4. Torres, N. V. & Santos, G. 2015 The (Mathematical) Modeling Process in Biosciences. *Front Genet*, **6**, 354. doi:10.3389/fgene.2015.00354.

5. Karr, J. R., Takahashi, K. & Funahashi, A. 2015 The principles of whole-cell modeling. *Curr Opin Microbiol*, **27**, 18–24. doi:10.1016/j.mib.2015.06.004.
6. Oreskes, N., Shrader-Frechette, K. & Belitz, K. 1994 Verification, validation, and confirmation of numerical models in the Earth sciences. *Science*, **263**(5147), 641–646. doi:10.1126/science.263.5147.641.
7. De Boer, R. J. 2012 Which of our modeling predictions are robust? *PLoS Comput Biol*, **8**(7), e1002593. doi:10.1371/journal.pcbi.1002593.
8. Chamberlin, T. C. 1890 The Method of Multiple Working Hypotheses: With this method the dangers of parental affection for a favorite theory can be circumvented. *Science*, **15**, 92–96.
9. Platt, J. R. 1964 Strong Inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, **146**(3642), 347–353. doi:10.1126/science.146.3642.347.
10. Hilborn, R. & Mangel, M. 1997 *The ecological detective: confronting models with data*. Princeton University Press, Princeton NJ.
11. Kirk, P. D. W., Babbie, A. C. & Stumpf, M. P. H. 2015 SYSTEMS BIOLOGY. Systems biology (un)certainities. *Science*, **350**(6259), 386–388. doi:10.1126/science.aac9505.
12. Popper, K. 2002 *The Logic of Scientific Discovery*. Routledge Classics. Taylor & Francis.
13. Noecker, C., Schaefer, K., Zaccheo, K., Yang, Y., Day, J. & Ganusov, V. V. 2015 Simple mathematical models do not accurately predict early SIV dynamics. *Viruses*, **7**, 1189–1217.